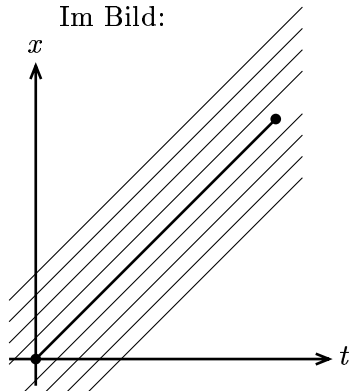


**(8.0) Einführung:** siehe die Aufgabe von Blatt 7. Im ganzen Kapitel setzen wir  $L \in C^3$  voraus.

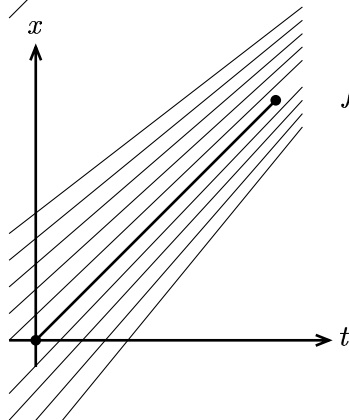
**(8.1) Paradigma:** Um zu zeigen, daß die Lösung  $x_*(t) = t$  der Eulergleichung des VP  $I[x] := \int_0^1 \dot{x}^2(t) dt$ ,  $x(0) = 0$ ,  $x(1) = 1$  wirklich ein Minimum ist, können wir wie folgt argumentieren (wir vergessen einmal, daß  $I[\cdot]$  konvex ist, um ein allgemeiner brauchbares Argument zu studieren):

$$\begin{aligned}
 I[x] &= \int_0^1 \dot{x}(t)^2 dt = \int_0^1 (\dot{x}(t) - 1)^2 dt & + & \int_0^1 (2\dot{x}(t) - 1) dt \\
 &= \int_0^1 (\dot{x}(t) - 1)^2 dt & + & \left[ 2x(t) - t \right]_0^1 \\
 &\text{Steigung 1 "gratis"} & + & \text{wegunabhängiger "Pauschalpreis" 1} \\
 &\text{andere Steigung kostet extra} & & (= I[x_*])
 \end{aligned}$$



Hier zeigen die dünnen Linien jeweils die "Gratis-Richtung" an; diese Linien sind übrigens selber Lösungen der Euler-Gleichung.

Es geht aber auch ganz anders, und doch nach demselben Prinzip:



$$\begin{aligned}
 \int_0^1 \dot{x}(t)^2 dt &= \int_0^1 \left( \dot{x}(t) - \frac{2-x(t)}{2-t} \right)^2 dt - \int_0^1 \frac{d}{dt} \left( \frac{(2-x(t))^2}{2-t} \right) dt \\
 &\text{Ortsabhängige "Gratis-Steigung"} & + & \text{wegunabhängiger "Pauschalpreis" 1}
 \end{aligned}$$

Hier definieren die Extremalen  $x(t; a) = a + (1 - \frac{a}{2})t$  die erwähnte Gratisrichtung. Durch Elimination von  $a$  folgt in der Tat  $\dot{x} = \frac{2-x}{2-t}$  für diese Extremalenschar. Hier wie im vorherigen Fall definieren die Gratisrichtungen Minimale, da der andere Beitrag zu  $I$  ein wegunabhängiger Pauschalpreis ist.

**(8.2) Dieses Verfahren geht allgemein:**

Es sei  $(t, x) \mapsto \psi(t, x)$ ,  $\mathbb{R} \times \mathbb{R}^n \supset \mathcal{G} \rightarrow \mathbb{R}^n$  ein in einer Umgebung  $\mathcal{G}$  einer Extremalen  $x_*$  definiertes Richtungsfeld. (Umgebung versteht sich hier also nicht in einem Funktionenraum, sondern in dem Raum  $\mathbb{R} \times \mathbb{R}^n$ , in dem der Graph der Funktion  $x(\cdot)$  liegt.) Dabei mögen die Integralkurven dieses Richtungsfeldes, d.h. die Lösungskurven der Differentialgleichung  $\dot{x}(t) = \psi(t, x(t))$ , selber Extremale sein, und  $x_*$  sei eine davon. Daß die Integralkurven selber Extremale sind, ist sicher notwendig, damit die Richtungen  $\psi(t, x)$  "Gratisrichtungen" sind und somit Minimale definieren.

Unser Ziel wird also sein, ein Vektorfeld  $\psi(t, x)$  zu bestimmen, und dazu eine skalarwertige Funktion  $S(t, x)$ , so daß wir für beliebige Kurven  $x(\cdot)$  schreiben können:

$$\begin{aligned}
 L(t, x(t), \dot{x}(t)) &= \tilde{L}(t, x(t), \dot{x}(t)) + \frac{d}{dt} S(t, x(t)) \quad , \text{ also:} \\
 L(t, x, \dot{x}) &= \tilde{L}(t, x, \dot{x}) + S_t(t, x) + \dot{x} S_x(t, x) \quad \text{für alle } (t, x, \dot{x}) ,
 \end{aligned}$$

wobei für alle  $(t, x)$  die Funktion  $\tilde{L}(t, x, \cdot)$  ein absolutes Minimum bei  $\psi(t, x)$  habe (und zwar mit Wert 0 daselbst). Notwendig hierfür ist laut Analysis 1:

$$\begin{aligned}
 L_{\dot{x}}(t, x, \psi(t, x)) &= S_x(t, x) \\
 L(t, x, \psi(t, x)) - \psi(t, x) L_{\dot{x}}(t, x, \psi(t, x)) &= S_t(t, x)
 \end{aligned} \tag{1}$$

notwendig, daß die Integrabilitätsbedingungen  $\partial_{x_j} a_i = \partial_{x_i} a_j$  und  $\partial_{x_j} b = \partial_t a_j$  erfüllt sind. Dies ist auch hinreichend, wenn das Definitionsgebiet der Funktionen  $a_i$  und  $b$  einfach zusammenhängend ist, und diese topologische Voraussetzung können wir an die Umgebung der Extremalen  $x_*$ , in der wir die Funktionen  $\psi$ ,  $S$  konstruieren wollen, leicht machen. Im Falle  $n = 1$  (skalarwertige Funktionen  $x(\cdot)$ ) tritt nur die Bedingung  $\partial_{x_j} b = \partial_t a_j$  auf, und es ist leicht, nachzurechnen, daß diese Bedingung genau bedeutet, daß die Integralkurven von  $\psi$  (d.h. die Lösungen von  $\dot{x}(t) = \psi(t, x(t))$ ) Extremale sind:

Beim Beweis müssen wir partielle Ableitungen  $\partial/\partial t$ , die sich auf Funktionen im  $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$  der drei Variablen  $(t, x, \dot{x})$  beziehen, unterscheiden von partiellen Ableitungen  $d/\partial t$ , die sich auf Funktionen im  $\mathbb{R} \times \mathbb{R}^n$  der zwei Variablen  $(t, x)$  beziehen, nachdem wir  $\dot{x} = \psi(t, x)$  eingesetzt haben. Ferner die totale Ableitung  $d/dt$ , die sich auf eine Funktion nur von  $t$  bezieht, nachdem auch für  $x$  eine konkrete Kurve  $t \mapsto x(t)$  eingesetzt wurde. Zur Verdeutlichung werden wir alle Terme, auf die sich die partielle Ableitung bezieht, mit Pfeilen markieren. Analoges gilt für partielle Ableitungen  $\partial/\partial x_i$  versus  $d/\partial x_i$ .

Die Integrabilitätsbedingung  $\partial_t a_i = \partial_{x_i} b$  lautet also (in Komponentenschreibweise)

$$\begin{aligned} \underbrace{\frac{d}{dt} L_{\dot{x}_i}(t, x, \psi(t, x))}_{\uparrow} &= \underbrace{\frac{d}{\partial x_i} \left( L(t, x, \psi(t, x)) - \sum_j \psi_j(t, x) L_{\dot{x}_j}(t, x, \psi(t, x)) \right)}_{\uparrow \uparrow \uparrow \uparrow \uparrow} \\ \underbrace{\frac{d}{dt} L_{\dot{x}_i}(t, x, \psi(t, x))}_{\uparrow} + \sum_j \psi_j(t, x) \underbrace{\frac{d}{\partial x_i} L_{\dot{x}_j}(t, x, \psi(t, x))}_{\uparrow} &= \\ &= \underbrace{\frac{d}{\partial x_i} L(t, x, \psi(t, x))}_{\uparrow} - \sum_j \underbrace{\left( \frac{\partial}{\partial x_i} \psi_j(t, x) \right)}_{\uparrow} L_{\dot{x}_j}(t, x, \psi(t, x)) \end{aligned}$$

als Gleichung von Funktionen in  $(t, x)$ . Auf dem Graphen einer Integralkurve  $t \mapsto x(t)$  des Feldes  $\psi$  ausgewertet ( $\dot{x}(t) = \psi(t, x(t))$ ) ergibt das

$$\frac{d}{dt} L_{\dot{x}_i}(t, x(t), \psi(t, x(t))) = L_{x_i}(t, x(t), \psi(t, x(t))) ,$$

das ist genau die Bedingung, daß Integralkurven die Eulergleichung erfüllen.

Für  $n = 1$  müssen wir also nur eine Umgebung des extremalen Segments  $x_*$  mit weiteren Extremalen fasn können (fasern heißt: durch jeden Punkt geht genau eine Extremale), und dann gibt uns die Ableitung der durch  $(t, x)$  gehenden Extremalen in diesem Punkt den Wert  $\psi(t, x)$  an. Für  $n > 1$  genügt nicht irgendeine Faserung durch Extremale, sondern wir müssen noch weitere Integrabilitätsbedingungen erfüllen. Daß auch dies unter Voraussetzungen wie in (8.5) möglich sein wird, werden wir später sehen. Jedenfalls haben wir das folgende programmatische Lemma schon fast bewiesen:

**(8.3) Programmatisches Lemma:** Falls es möglich ist, in einer Umgebung einer Extremalen  $t \mapsto x_*(t)$ ,  $[t_0, t_1] \rightarrow \mathbb{R}^n$  ein Vektorfeld  $\psi : (t, x) \mapsto \psi(t, x) \in \mathbb{R}^n$  derart zu definieren, daß  $\dot{x}_*(t) = \psi(t, x_*(t))$ , und daß für

$$a_i(t, x) := L_{\dot{x}_i}(t, x, \psi(t, x)) \tag{8.3.1}$$

$$b(t, x) := L(t, x, \psi(t, x)) - \sum_j \psi_j(t, x) L_{\dot{x}_j}(t, x, \psi(t, x)) \tag{8.3.2}$$

die Integrabilitätsbedingungen

$$\partial_{x_j} a_i = \partial_{x_i} a_j \tag{8.3.3}$$

$$\partial_{x_j} b = \partial_t a_j \tag{8.3.4}$$

erfüllt sind, dann kann man mit einer Stammfunktion  $S$  (d.h.  $\partial_{x_j} S = a_j$ ,  $\partial_t S = b$ ) schreiben:

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt = \int_{t_0}^{t_1} \mathfrak{E}(t, x(t), \psi(t, x(t)), \dot{x}(t)) dt + \left[ S(t, x(t)) \right]_{t_0}^{t_1} ,$$

$$= (\dot{X} - \psi)^2 \int_0^1 L_{\dot{x}\dot{x}}(t, X, (1-s)\psi + s\dot{X}) ds .$$

**Beweis:**

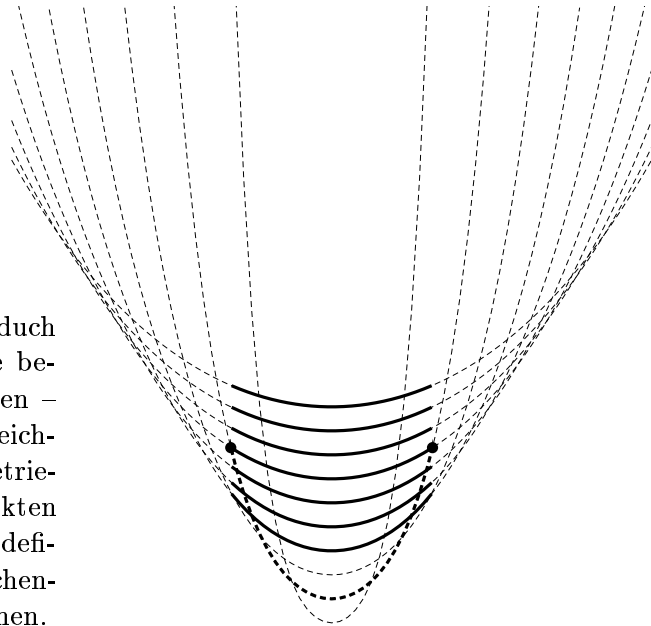
$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt = \int_{t_0}^{t_1} \underbrace{(L(t, x, \dot{x}) - a\dot{x} - b)}_{\mathfrak{E}(\dots)} dt + \int_{t_0}^{t_1} \underbrace{(a\dot{x} + b)}_{\frac{d}{dt}S(t, x(t))} dt .$$

**Bemerkungen:**

- (1) Das wegunabhängige Integral  $S = \int (L + (\dot{x} - \psi)L_{\dot{x}}(t, x, \psi)) dt$  heißt *Hilberts invariantes Integral*.
- (2) Das Lemma ist relevant, weil wir in der Lage sein werden, die Existenz von  $\psi$  unter leicht prüfbaren Voraussetzungen zu zeigen.
- (3) Konvexität von  $L$  *allein in der Variablen  $\dot{x}$*  garantiert  $\mathfrak{E} \geq 0$  und damit unter den Voraussetzungen des Lemmas Minimalität nach dem Argument des Paradigmas (8.1).
- (4) Die wesentliche Obstruktion gegen die Konstruktion von  $\psi$  ist, daß sich benachbarte Extremale im Intervall  $[t_0, t_1]$  schneiden können. Da aber nach Aufgabe 27a (Blatt 7) der Abstand benachbarter Extremaler (in führender Ordnung) durch die Jacobigleichung bestimmt wird, sind *konjugierte Punkte* die infinitesimale Version dieser Obstruktion.

#### (8.4) Beispiel: Das Catenoid

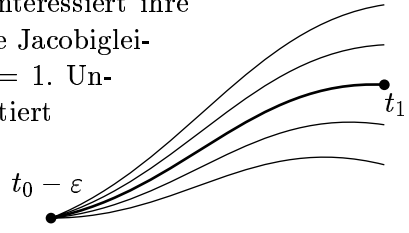
Hier ist die in Aufgaben 10 und 11 berechnete Schar von Extremalen gestrichelt eingezeichnet. Zwischen den markierten Punkten sei das flachere Segment betrachtet. Eine naheliegende Faserung einer Umgebung dieses Segments mit anderen extremalen Segmenten ist durch dicke Linien herausgehoben.



Anders die Situation beim steilen Segment durch die beiden Punkte: dieses wird durch die benachbarten Extremalensegmente geschnitten – und das gilt auch für jene (nicht eingezeichneten) Extremalensegmente, deren Symmetrieachsen  $x = x_0 \neq 0$  sind. In Schnittpunkten von Extremalen kann also  $\psi(t, x)$  nicht definiert werden, da dafür zwei sich widersprechende Werte der Kurvensteigung in Frage kämen.

**(8.5) Korollar:** Sei  $n = 1$ . Falls  $x_*$  auf  $[t_0, t_1]$  Extremale ist mit  $L_{\dot{x}\dot{x}}^* > 0$ , und falls es obendrein auf dem Segment keinen zu  $x_*(t_0)$  konjugierten Punkt gibt, dann ist  $x_*$  eine schwache Minimale. Falls zusätzlich gilt:  $\mathfrak{E}(t, x, \dot{x}, \dot{X}) \geq 0$  in einer Umgebung des Segments, d.h. für alle  $X$  aber für  $|x - x_*(t)| < \varepsilon$ ,  $|\dot{x} - \dot{x}_*(t)| < \varepsilon$ , dann ist  $x_*$  eine starke Minimale.

Betrachte Lösungen  $t \mapsto x(t; v)$  der Eulergleichung zu den Anfangsbedingungen  $x(t_0) = x_*(t_0)$ ,  $\dot{x}(t_0) = v$ ,  $v$  nahe bei  $\dot{x}_*(t_0)$ . Insbesondere interessiert ihre Abhängigkeit von  $v$ :  $\varphi(t) := \partial x(t; v) / \partial v$  löst bei  $v = \dot{x}_*(t_0)$  die Jacobigleichung um  $x_*$  mit den Anfangsbedingungen  $\varphi(t_0) = 0$ ,  $\dot{\varphi}(t_0) = 1$ . Unter der Annahme, daß es keine konjugierten Punkt gibt, garantiert der Satz über implizite Funktionen, daß die Extremalenschar  $x(t; v)$  für  $v$  in einer Umgebung von  $\dot{x}_*(t_0)$  keine Schnittpunkte außer  $(t_0, x_*(t_0))$  hat. Das ist noch nicht ganz, was wir wollen. Aber wenn es zu  $(t_0, x_*(t_0))$  keinen konjugierten Punkt auf  $[t_0, t_1]$  gibt, dann aus Stetigkeitsgründen auch nicht zu  $(t_0 - \varepsilon, x_*(t_0 - \varepsilon))$ , und wir wenden das obige Argument auf diesen Fall an. So können wir eine Umgebung des Segments  $x_*|_{[t_0, t_1]}$  mit Extremalen fasern.



Als eine weite ( $C^0$ -) Umgebung von  $x_*$  können wir die Gesamtheit all derjenigen Kurven nehmen, die in dem von der Extremalenschar gefaserten Bereich liegen; als eine enge ( $C^1$ -) Umgebung die Gesamtheit derjenigen Kurven, für die obendrein noch  $\|\dot{x} - \dot{x}_*\|_\infty < \varepsilon$  ist.

Wegen

$$I[x] = \int_{t_0}^{t_1} \mathfrak{E}(t, x(t), \psi(t, x(t)), \dot{x}(t)) dt + I[x_*]$$

folgt, daß  $I[x] \geq I[x_*]$  für alle Kurven, für die (punktweise)  $\mathfrak{E} \geq 0$  ist. Daher ist  $x_*$  starkes Minimum, wenn die Weierstraßbedingung gilt. Falls aber nur  $L_{\dot{x}\dot{x}}^* > 0$  gilt, so gibt es eine kleine Umgebung  $\{(t, x, \dot{x}) \mid t \in [t_0, t_1], |x - x_*(t)| < \varepsilon, |\dot{x} - \dot{x}_*(t)| < \varepsilon\} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ , auf der immer noch  $L_{\dot{x}\dot{x}} > 0$  gilt. Nach eventueller Verkleinerung von  $\varepsilon$  liegt diese in dem oben überdeckten Bereich. Dies definiert eine enge Umgebung von  $x_*$ , und in dieser gilt immer noch  $\mathfrak{E}(t, x(t), \psi(t, x(t)), \dot{x}(t)) \geq 0$ , mit Gleichheit nur für  $\dot{x}_* \equiv \dot{x}$ .

Als nächstes stellen wir uns der Aufgabe, ein geeignetes Extremalenfeld auch im Fall  $n > 1$  zu konstruieren:

**(8.6) Definition:** Ein Vektorfeld  $\psi(t, x)$  (oder auch die zugehörige Schar von Integralkurven) heißt *Mayerfeld* (zu einer Lagrangefunktion  $L$ ), wenn es die Integrabilitätsbedingungen (8.3.3), (8.3.4) erfüllt, von denen, wie gesagt, (8.3.4) allein gleichbedeutend damit ist, daß die Kurven ein *Extremalenfeld* bilden.

**(8.6a) Satz:** Ein Extremalenfeld zu festem Anfangspunkt ist automatisch ein Mayerfeld; genauer: Sei  $x(\cdot, v)$ ,  $v \in \mathcal{V} \subset \mathbb{R}^n$ ,  $\mathcal{V}$  offen die Schar der Lösungen zur Eulergleichung mit den Anfangsbedingungen  $x(t_0; v) = x_*$  fest,  $\dot{x}(t_0; v) = v$ . Für ein Intervall  $]t_0, t_1[$ , auf dem sich die Scharcurven nicht schneiden (d.h.  $x(t; v_1) = x(t; v_2) \wedge t \in ]t_0, t_1[ \implies v_1 = v_2$ ), ist durch  $\psi(t, x(t; v)) := \dot{x}(t; v)$  auf  $\mathcal{G} := x(]t_0, t_1[, \mathcal{V})$  ein Mayerfeld definiert. Für  $t_1$  hinreichend nahe an  $t_0$  ist die Injektivitätsbedingung an  $x(t; \cdot)$  erfüllt und die Matrix  $\partial x / \partial v$  invertierbar.

**Beweis:** Daß für  $|t_1 - t_0|$  klein die Injektivitätsbedingung erfüllt ist und  $\partial x / \partial v$  invertierbar ist, folgt wegen  $Dx(t; \cdot) = (t - t_0)\mathbf{1} + O(t - t_0)^2$  aus dem Satz über implizite Funktionen. Unter dieser Bedingung ist  $\psi$  auf  $\mathcal{G}$  definiert und  $C^1$ . Wir haben die Integrabilitätsbedingung (8.3.3) zu prüfen. Mit

$$p_i(t; v) := a_i(t, x(t; v)) = L_{\dot{x}_i}(t, x(t; v), \dot{x}(t; v))$$

lautet sie  $\partial a_i / \partial x_j = \partial a_j / \partial x_i$ . Wegen der Invertierbarkeit der Matrix  $\partial x / \partial v$  in  $\mathcal{G}$  ist sie äquivalent zu

$$\sum_{i,j} \frac{\partial a_i}{\partial x_j} \frac{\partial x_j}{\partial v_k} \frac{\partial x_i}{\partial v_r} = \sum_{i,j} \frac{\partial a_j}{\partial x_i} \frac{\partial x_j}{\partial v_k} \frac{\partial x_i}{\partial v_r}.$$

Nach der Kettenregel für  $p(t; v) = a(t, x(t; v))$  ist dies äquivalent zu

$$\sum_i \left( \frac{\partial p_i}{\partial v_k} \frac{\partial x_i}{\partial v_r} - \frac{\partial p_i}{\partial v_r} \frac{\partial x_i}{\partial v_k} \right) = 0. \quad (8.6.1)$$

Der Ausdruck auf der linken Seite heißt Lagrangeklammer der Kurvenschar  $x(t; \cdot)$ . Die Lagrangeklammer hängt bei *Extremalenscharen* allgemein nur vom Scharparameter ab, nicht aber von  $t$ ,

Wir zeigen jetzt die zeitliche Konstanz der Lagrangeklammer für Extremalenscharen:

$$\begin{aligned}
& \frac{\partial}{\partial t} \sum_i \left( \frac{\partial p_i}{\partial v_k} \frac{\partial x_i}{\partial v_r} - \frac{\partial p_i}{\partial v_r} \frac{\partial x_i}{\partial v_k} \right) = \sum_i \left( \left( \frac{\partial}{\partial v_k} \frac{\partial p_i}{\partial t} \right) \frac{\partial x_i}{\partial v_r} + \frac{\partial p_i}{\partial v_k} \frac{\partial \dot{x}_i}{\partial v_r} - (k \leftrightarrow r) \right) \\
& = \sum_i \left( \left( \frac{\partial}{\partial v_k} L_{x_i} \right) \frac{\partial x_i}{\partial v_r} + \left( \frac{\partial}{\partial v_k} L_{\dot{x}_i} \right) \frac{\partial \dot{x}_i}{\partial v_r} - (k \leftrightarrow r) \right) \\
& = \sum_{i,s} \left( L_{x_i x_s} \frac{\partial x_s}{\partial v_k} \frac{\partial x_i}{\partial v_r} + L_{x_i \dot{x}_s} \frac{\partial \dot{x}_s}{\partial v_k} \frac{\partial x_i}{\partial v_r} + L_{\dot{x}_i x_s} \frac{\partial x_s}{\partial v_k} \frac{\partial \dot{x}_i}{\partial v_r} + L_{\dot{x}_i \dot{x}_s} \frac{\partial \dot{x}_s}{\partial v_k} \frac{\partial \dot{x}_i}{\partial v_r} - (k \leftrightarrow r) \right) \\
& = 0
\end{aligned}$$

Damit ist der Satz bewiesen und wir haben

**(8.6b) Satz:** Korollar (8.5) gilt mit dem analogen Argument auch ohne die Voraussetzung “ $n = 1$ ”.

**Damit haben wir den hinreichenden Teil der Bedingungen auf dem zweiten Übersichtsblatt über verschiedene Minimalitätsbegriffe bewiesen. Der notwendige Teil folgt aus den Übungen von Blatt 7 und 8.**

**(8.7) Satz (mit Def.):** Unter der Voraussetzung  $L_{\dot{x}\dot{x}} > 0$  (d.h. positiv definit) ist die Abbildung  $\dot{x} \mapsto L_{\dot{x}}(t, x, \dot{x}) =: p$  injektiv, und die Eulergleichung  $\frac{d}{dt} L_{\dot{x}}(t, x(t), \dot{x}(t)) = L(t, x(t), \dot{x}(t))$  läßt sich als System der sog. *Hamilton-Gleichungen*

$$\begin{aligned}
\frac{d}{dt} p(t) &= -H_x(t, x(t), p(t)) \\
\frac{d}{dt} x(t) &= H_p(t, x(t), p(t))
\end{aligned} \tag{8.7.1}$$

schreiben, wobei die sog. *Hamiltonfunktion*  $H$  durch

$$\begin{aligned}
H(t, x, p) &= \dot{x} L_{\dot{x}}(t, x, \dot{x}) - L(t, x, \dot{x}) \\
p &= L_{\dot{x}}(t, x, \dot{x})
\end{aligned} \tag{8.7.2}$$

definiert ist. Diese Transformation heißt *Legendre-Transformation*.

**Beweis:** Die Injektivität der Legendretransformation  $p \leftrightarrow \dot{x}$  hatten wir schon in Punkt (2.13) der Vorlesung. Diese Transformation ist in beide Richtungen  $C^1$ , wenn  $L \in C^2$  (Satz über implizite Funktionen). Wir betrachten die Legendre-Transformation (kinematisch, d.h. mit  $(t, x, \dot{x})$  als unabhängigen Variablen). Durch Differenzieren von

$$H(t, x, L_{\dot{x}}(t, x, \dot{x})) = \dot{x} L_{\dot{x}}(t, x, \dot{x}) - L(t, x, \dot{x})$$

nach  $\dot{x}_i$  bzw.  $x_i$  folgt:

$$\begin{aligned}
\sum_j H_{p_j} L_{\dot{x}_j \dot{x}_i} &= L_{\dot{x}_i} + \sum_j \dot{x}_j L_{\dot{x}_j \dot{x}_i} - L_{\dot{x}_i} & \text{also} & \quad H_p = \dot{x} \\
H_{x_i} + \sum_j H_{p_j} L_{\dot{x}_j x_i} &= \sum_j \dot{x}_j L_{\dot{x}_j x_i} - L_{x_i} & & \quad H_x = -L_x
\end{aligned}$$

Daraus ergibt sich (dynamisch, d.h. jetzt  $x(t)$  eingesetzt):

$$\begin{aligned}
\frac{d}{dt} p(t) &\stackrel{\text{Def. } p}{=} \frac{d}{dt} L_{\dot{x}}(t, x(t), \dot{x}(t)) \stackrel{\text{Eulerglg.}}{=} L_x(t, x(t), \dot{x}(t)) \stackrel{\text{KooTrf.}}{=} -H_x(t, x(t), p(t)) \\
\frac{d}{dt} x(t) &= \dot{x}(t) \stackrel{\text{KooTrf.}}{=} H_p(t, x(t), p(t))
\end{aligned}$$

Übrigens folgt analog, daß  $L_{\dot{x}\dot{x}}$  und  $H_{pp}$  zueinander inverse Matrizen sind, insbesondere ist also  $H_{pp}$  positiv definit.

**Bemerkung:** Der Funktion  $H$  sind wir schon in (3.2) begegnet, wir hatten nämlich für  $L_t = 0$  die Eulergleichungen als  $\frac{d}{dt} H = 0$  geschrieben. Erinnerung aus der Mechanik:  $H$  ist dort die Energie,  $\dot{x}$  die Geschwindigkeit,  $p$  der Impuls.

In Hamiltonscher Formulierung schreiben sich die sog. *Fundamentalgleichungen der Variationsrechnung* (8.3.1) als

$$\begin{aligned}
\partial_{x_i} S &= L_{\dot{x}_i}(t, x, \psi(t, x)) & = p_i(t, x) \\
\partial_t S &= L(t, x, \psi(t, x)) - \psi(t, x) L_{\dot{x}}(t, x, \psi(t, x)) & = -H(t, x, p(t, x))
\end{aligned}$$

$$S_t(t, x) + H(t, x, S_x(t, x)) = 0 \quad (8.7.3)$$

$$\psi(t, x) = H_p(t, x, S_x(t, x)) \quad (8.7.4)$$

Wir haben somit

**(8.8) Korollar:** Die Voraussetzung von Lemma (8.3), d.h. die Existenz eines Mayerfeldes  $\psi$  in einer Umgebung einer Extremalen, ist äquivalent mit der Lösbarkeit der Gleichung (8.7.3), wobei  $H$  die zur Lagrangefunktion  $L$  gehörige Hamiltonfunktion ist.

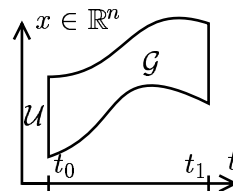
**(8.8a) Definition:** Die Gleichung (8.7.3) heißt *Hamilton–Jacobi–Gleichung*.

Ihre Lösbarkeit ist äquivalent zur Existenz eines Mayerfeldes, und im benötigten Rahmen haben wir diese in (8.6a) bereits gezeigt; dem Zusammenhang zwischen den Hamilton–Gleichungen (8.7.1) und der Hamilton–Jacobi–Gleichung liegt ein allgemeines Prinzip zugrunde, das wir nachfolgend noch ausführen. Partielle Differentialgleichungen erster Ordnung (auch eine Auswahl von solchen höherer Ordnung wie z.B. die Wellengleichung) lassen sich mit der *Methode der Charakteristiken* auf gewöhnliche Differentialgleichungen zurückführen.

**(8.9) Satz:** Für  $H \in C^2$ ,  $H_{pp} > 0$  (d.h. positiv definit),  $\mathcal{U}$  ein beschränktes Gebiet im  $\mathbb{R}^n$  und eine gegebene Funktion  $g \in C^3(\bar{\mathcal{U}} \rightarrow \mathbb{R})$  (welche vorgegebene Anfangswerte beschreibe) läßt sich eine Lösung  $S \in C^3$  des sog. *Cauchyproblems* für die Hamilton–Jacobi–Gleichung

$$S_t + H(t, x, S_x) = 0 \quad \text{für } (t, x) \in \mathcal{G}; \quad \mathcal{G} \text{ geeignet wählbar}$$

$$S(t_0, x) = g(x) \quad \text{für } x \in \mathcal{U}; \quad \mathcal{U} \text{ gegeben}$$



(8.9.1)

wie folgt konstruieren: Betrachte das Anfangswertproblem für das System gewöhnlicher Differentialgleichungen

$$\begin{aligned} \dot{x}(t) &= H_y(t, x(t), y(t)) & x(t_0) &=: u \in \mathcal{U} \\ \dot{y}(t) &= -H_x(t, x(t), y(t)) & y(t_0) &= \nabla g(u) \\ \dot{z}(t) &= -H(t, x(t), y(t)) + y(t)H_y(t, x(t), y(t)) & z(t_0) &= g(u) \end{aligned} \quad (8.9.2)$$

Wenn die Abbildung  $(t, u) \mapsto (t, x)$ ,  $[t_0, t_1] \times \mathcal{U} \rightarrow \mathcal{G}$  (die Projektion der Flußabbildung auf die  $x$ -Komponenten) bijektiv ist, dann hat das Cauchy–Problem in  $\mathcal{G}$  eine eindeutige Lösung; wenn die Abbildung aber nicht injektiv ist, hat das Problem in  $\mathcal{G}$  keine Lösung. Im ersten Fall wird zu gegebenem  $(t, x) \in \mathcal{G}$  eindeutig der Anfangswert  $u$  bestimmt, für welchen die Lösung durch  $(t, x)$  geht, und mit der Lösung zu diesem Anfangswert definiert man  $z(t) =: S(t, x)$ .

**Vorbemerkung zum Beweis:** Die Methode der Charakteristiken besteht darin, daß man im Gebiet, auf dem die zu findende Lösung der partiellen Differentialgleichung definiert sein soll, Kurven sucht (sog. *charakteristische Kurven*), längs derer sich die PDgl zu einer gewöhnlichen Differentialgleichung reduziert (nicht für jede PDgl gibt es solche Kurven). Zum Beispiel sind für die PDgl  $u_x(t, x) = 2u_t(t, x)$  die Kurven  $x = a - t/2$  solche Kurven; denn mit  $w(t) := u(t, a - t/2)$  reduziert sich die PDgl auf die gewDgl  $w'(t) = 0$ . Diese Lösungstechnik findet im Beweis Anwendung, und die Dgl für die charakteristischen Kurven sind gerade die Hamiltonschen Gleichungen.

**Beweis:** Wir nehmen zunächst an, wir hätten eine Lösung  $S(t, x)$  von (8.9.1) und definieren auf irgendeiner Kurve  $(t, x(t))$  in ihrem Definitionsgebiet (und abhängig von der gewählten Kurve)  $z(t) := S(t, x(t))$ . Zunächst die Motivation:

Wir wollen die Kurve so bestimmen, daß sich die PDgl (8.7.3) als eine gewDgl für  $z$  schreiben läßt. Wir finden

$$\dot{z}(t) = S_t(t, x(t)) + \sum_j S_{x_j}(t, x(t)) \dot{x}_j(t) = -H(t, x(t), S_x(t, x(t))) + \sum_j S_{x_j}(t, x(t)) \dot{x}_j(t)$$

Es ist also noch nicht gelungen,  $S$  zu eliminieren:  $S_x$  auf der Kurve ist nicht durch  $S$  auf der Kurve bestimmt; also definieren wir auch noch  $y(t) := S_x(t, x(t))$  und versuchen, die Kurve so

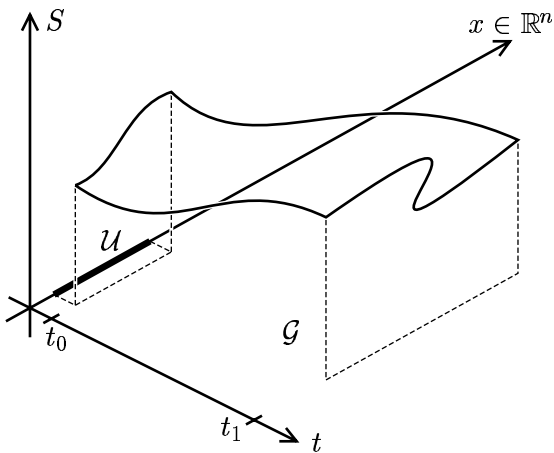
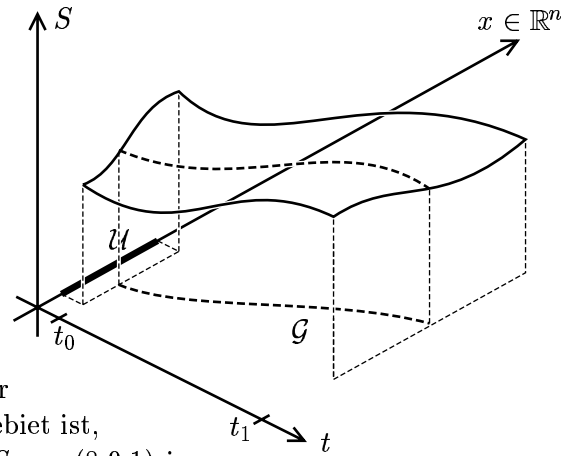
Ableitung von (8.9.3), nämlich  $S_{tx_k} + H_{x_k}(t, x, S_x(t, x)) + \sum_j H_{y_j}(t, x, S_x(t, x)) S_{x_j x_k}(t, x) = 0$ , daß

$$\begin{aligned} \dot{y}_k(t) &= S_{x_k t}(t, x(t)) + \sum_j S_{x_k x_j}(t, x(t)) \dot{x}_j(t) \\ &= -H_{x_k}(t, x(t), y(t)) - \sum_j S_{x_k x_j}(t, x(t)) (H_{y_j}(t, x(t), y(t)) - \dot{x}_j(t)) \end{aligned}$$

Wir verlangen also, um nicht noch mehr Ableitungen von  $S$  in unser System von gewDgl aufnehmen zu müssen, daß die große Klammer verschwindet und erhalten so eine Dgl. für die charakterisitischen Kurven, nämlich genau das System (8.9.2).

Zusammenfassend betrachten wir also eine Lösung  $S$  von (8.9.1) auf einem Kurvenstück  $x(t)$  im Definitionsbereich von  $S$ , welches durch die Hamiltonschen Gleichungen  $\dot{x} = -H_y$ ,  $\dot{y} = -H_x$  bestimmt wird. Auf einem solchen Kurvenstück genügt dann  $S(t, x(t)) =: z(t)$  der gewDgl  $\dot{z} = -H + yH_y$ , die aus der eingangs berechneten  $\dot{z} = -H + S_x \dot{x}$  durch Einsetzen folgt.

Wenn wir auf  $\{t_0\} \times \mathcal{U}$  Werte für  $S$  vorgeben, sind dort natürlich die Werte von  $S_x$  durch Differentiation ebenfalls bestimmt; so folgen die Anfangsbedingungen in (8.9.2). Wenn also  $\mathcal{G}$  das durch Lösen der Hamiltongleichungen (8.9.2) mit  $u \in \mathcal{U}$  bestimmte Gebiet ist, dann haben wir gezeigt, daß eine eventuelle Lösung  $S$  von (8.9.1) in diesem Gebiet durch die  $\dot{z}$ -Gleichung von (8.9.2) eindeutig bestimmt ist.



Umgekehrt ist aber nicht sicher, daß jede Lösung von (8.9.2) zu einer Lösung von (8.9.1) führt. Es könnte ja sein, daß Lösungen  $(t, x, y, z)$  zu verschiedenen Anfangswerten in den Koordinaten  $(t, x)$  übereinstimmen. Im Bild würde das bedeuten, daß die aus den Lösungskurven von (8.9.2) bestehende Fläche nicht der *Graph einer Funktion* ist, sondern "Falten schlägt". Dann hat (8.9.1) keine Lösung in  $\mathcal{G}$ .

Wenn wir aber die Injektivitätsannahme des Satzes machen und  $S(\hat{t}, \hat{x})$  für  $(\hat{t}, \hat{x}) \in \mathcal{G}$  als den Wert  $z(\hat{t})$  der einzigen Lösung von (8.9.2) definieren, die  $x(\hat{t}) = \hat{x}$  erfüllt, dann haben wir als

Voraussetzung die Existenz von Funktionen  $x(t; u)$ ,  $y(t; u)$ ,  $z(t; u)$ , die dem AWP (8.9.2) genügen sowie der Gleichung

$$z(t; u) = S(t, x(t; u)) \tag{8.9.3}$$

für eine in  $\mathcal{G}$  definierte Funktion  $S$  genügen. Da  $\nabla g \in C^2$ , ist auch die Abhängigkeit  $x \leftrightarrow u$  ein  $C^2$ -Diffeomorphismus, und es folgt  $S \in C^2$ . Wir wollen zeigen, daß

$$S_t(t, x(t; u)) = -H(t, x(t; u), y(t; u)) \quad , \quad S_x(t, x(t; u)) = y(t; u)$$

gelten. Dann folgt sofort die PDgl (8.9.1), sowie  $S \in C^3$  wegen  $x, y \in C^2$ .

Dazu müssen wir natürlich die Anfangsbedingungen in (8.9.2) mitverwenden, weil ohne sie die Behauptung nicht einmal für  $t = t_0$  gelten würde. Wir differenzieren (8.9.3) nach der Zeit und erhalten mittels (8.9.2)

$$-H + yH_y = S_t + S_x H_y$$

bzw.

$$S_t(t, x(t; u)) + H(t, x(t; u), y(t; u)) = H_y(t, x(t; u), y(t; u)) (y(t; u) - S_x(t, x(t; u))) \tag{8.9.4}$$

weiter, um nachherend  $S_t$  bzw. die Zeitableitung in  $S_x$  zu eliminieren. Es folgt nämlich durch Differentiation nach  $u$  aus (8.9.4):

$$\sum_j \left( S_{tx_j} + H_{x_j} \right) \frac{\partial x_j}{\partial u_k} + H_{y_j} \frac{\partial y_j}{\partial u_k} = \sum_j \left\{ \frac{\partial H_{y_j}(\dots)}{\partial u_k} (y_j - S_{x_j}) + H_{y_j} \left( \frac{\partial y_j}{\partial u_k} - \sum_i S_{x_j x_i} \frac{\partial x_i}{\partial u_k} \right) \right\} \quad (8.9.5)$$

und somit

$$\sum_j \left( S_{tx_j} + H_{x_j} + \sum_i H_{y_i} S_{x_i x_j} \right) \frac{\partial x_j}{\partial u_k} = \sum_j \frac{\partial H_{y_j}(\dots)}{\partial u_k} (y_j - S_{x_j}) \quad (8.9.6)$$

Da nun  $S_x(t, x(t; u)) = y(t; u)$  wegen der Anfangsbedingungen jedenfalls für  $t = t_0$  gilt, bleibt noch die Ableitung dieser Gleichung nach  $t$  zu beweisen. Es ist aber

$$\frac{\partial}{\partial t} \left( S_x(t, x(t; u)) - y(t; u) \right) = S_{xt} + S_{xx}(t, x(t; u)) H_y + H_x$$

und daher mit (8.9.6)

$$\frac{\partial x}{\partial u} \cdot \frac{\partial}{\partial t} (S_x - y) = - \frac{\partial H_y}{\partial u} \cdot (S_x - y)$$

Wegen des Eindeutigkeitsatzes (und weil die Matrix  $\partial x / \partial u$  invertierbar ist) folgt  $S_x - y = 0$  für alle Zeit.

**(8.10) Interpretation:** Dem Fermatsche Prinzip zufolge bewegt sich Licht (zB. in inhomogenem Medium) entlang Kurven, die die Laufzeit (lokal) minimieren. Dies ist die Sichtweise der *Strahlenoptik*. Demgegenüber versteht die *Wellenoptik* Licht als elektromagnetische Welle. [Diese Dualität der Sichtweise, von der hier zu reden ist, ist eine Dualität der Interpretation klassischer Physik, und sie beruht nicht auf dem Welle-Korpuskel-Dualismus der Quantenmechanik.] Bei einer Welle sind die markanten geometrischen Objekte Wellenfronten, die man sich als Hyperflächen vorstellen kann, auf denen die Welle Knotenlinien hat (und die sich fortbewegen). Bei fester Frequenz, d.h. fester Anzahl von Schwingungen pro Zeiteinheit) sind solche Wellenfronten genau Flächen konstanter Laufzeit.

Den Lichtstrahlen entsprechen die Extremalen. Allerdings sollte man sich unter einem Extremalenfeld mehr vorstellen als nur die Lichtstrahlen aus einem Lichtblitz. Für  $I = \int \dot{x}^2$  wäre ja  $x(t) = ct$  mit Scharparameter  $c$  für  $t > 0$  ein Extremalenfeld, und es kommen alle „Lichtgeschwindigkeiten“  $c$  vor. Das Fermat-Prinzip spezifiziert nicht die Lichtgeschwindigkeit, und das zu Recht, denn es gilt ja in verschiedenen Medien. Da  $\int S$  ja das Funktional längs Extremalen des Feldes liefert, ist dies also die Laufzeit des Lichts (oder entspricht ihr zumindest, wenn wir mit  $I = \int \dot{x}^2$  statt  $I = \int |\dot{x}|$  arbeiten), und Flächen mit konstantem  $S$  entsprechen Wellenfronten; diese stehen senkrecht auf den Lichtstrahlen (im räumlichen Diagramm für Licht einer Quelle; nicht im Raum-Zeit-Diagramm).

Aber nicht jede Ansammlung von Lichtstrahlen (Extremalenfeld) zerfällt in Scharen, die zu ein und derselben Quelle gehören können. Das entspricht der Integrabilitätsbedingung für Mayerfelder.